Proceedings of the Canadian Society for Mechanical Engineering International Congress 2022
CSME Congress 2022
June 05-08, 2022, Edmonton, AB, Canada

# VISUAL CONTROL OF A ROBOTIC MANIPULATOR FOR TRACKING OF AN UNCOOPERATIVE PLATFORM

Muhammad Salman Chaudhry
*Lassonde School of Engineering*
*York University*
*Toronto, Canada*
*salman01@yorku.ca*

Aleksander Czekanski*
*Lassonde School of Engineering*
*York University*
*Toronto, Canada*
*alex.czekanski@lassonde.yorku.ca*

***Abstract*—This work presents a methodology for the visual detection and robotic tracking of an uncooperative target. This article presents a closed-loop visual feedback-based control strategy combined with the Kalman filtering technique for autonomous tracking of the target object. The framework is based on the independent implementation of visual detection and tracking control loop to enhance the detection and tracking precision of the object. The overall hardware setup consists of low-cost off-the-shelf hardware for marker-based visual detection. Through the adopted approach, the performance of the developed tracker and control is evaluated.**

## I. INTRODUCTION

Robotic manipulators are great for industrial and specialized applications that are humanely complex or sometimes even dangerous to perform. An example of this is the surgical robot PUMA [1] which was introduced in the 1980s. Robot-assisted surgery has since evolved substantially and the FDA-approved da Vinci surgical robot was introduced in the 2000s. Many modern surgical robots are developed around this system[2]. Today there are numerous competitive systems commercially being developed and used[3][4][5][6][7]. Robotics has proved extremely reliable in assisting and performing complex surgical operations. With the development of techniques in computer vision, it is now possible to integrate precise look and move type of visual servo architecture for complex applications like robotic surgery and in-situ fabrication. The visual sensors serve as the eyes for the robot controller that updates the status of the target, region of the target, and self-localization. Robotic tasks of interest consist of tracking an uncooperative object and then traversing that object to perform some operations, given that the geometry of the target is known in advance.

This work aims to develop a vision-based tracking and control scheme for a robotic manipulator to perform a track and traverse operation. The objective of the 6 dof robotic manipulator is to perform a target following an operation on an uncooperative platform with a high degree of accuracy (0.5 mm). The particular challenge of this study is to embed the two motions (track and traverse) into one motion command constantly sent out in the form of GCODE. The system is designed and built using low cost, off the shelf, ready to use hardware. The demonstration of the proposed methodology in this work is limited to 2 dof motion in the platform. Meaning the target is free to move in 2 dof, while the robotic arm performs two operations in parallel, tracks, and traverses through the platform. The experimental results show the effectiveness and robustness of the presented methodology by successfully performing these two operations simultaneously. The proposed approach presented for tracking and traverse control on uncooperative mobile platform is intended for new and novel applications like in-situ 3D printing, and robotic surgery.

## II. LITERATURE REVIEW

In complex robotic applications like surgical operations, vision system is used mainly for tool detection [8] localization [6][9] pose estimation [10][11] and tracking [12][13]. Various techniques that can be found in literature attributed to such problems ranging from machine learning to stereo imaging. We have mentioned a few recent examples for reference. In machine learning, convolution neural networks can be employed for the segmentation of surgical instrumentation, localization and prediction of the kinematic pose [14][8][15][7]. For example, in [6], 3d localization was performed by combing image information, deep learning, and bone-tracking data to estimate the camera pose relative to the bone markers. Moreover, deep learning can also be applied for efficient tracking of surgical instrumentation by detection in minimally invasive surgery [13]. Kalman filter is one of the most widely used tracking filters and its application in robotics is intended for object tracking and sensory data fusion [16][17][18]. Kalman filter is an optimal estimation algorithm that produces state estimates in the presence of noisy measurement. Another example of an optimal state estimation filter used for tracking is the particle filter[9][19]. Simultaneous localization and mapping (SLAM) approaches have also been proposed for the estimation of camera motion from a sequence of images[20]. Other techniques used for visual tracking and detection in robotics applications include the use of stereo imaging[21],

template-based matching [22][23], optical flow [24][25]and mean-shift algorithm [26]. To efficiently track an uncooperative object for this study Kalman filter for state estimation and tracking is applied. An advantage of employing Kalman filter is the separate measurement update and estimation step. This enables the fusion of multiple sensory data and optimal estimation in the presence of process noise. As discussed later in this study the visual detection loop is independent of the tracking and control loop. This allows the application of computationally heavy and precise image detection schemes. The methodology for this is explained in the next sections.

## III. Methodology

### A. Visual Servoing Control

Visual Servoing system can typically be classified by camera configuration, either the camera is mounted on the robotic manipulator called the eye-in-hand configuration or the camera is fixed in the work-space, called the eye-to-hand configuration. In this case, the image coordinates of the target are independent of the robotic motion. Visual Servoing systems can also typically be classified by control type. $Pose-based$ visual servoing(PBVS) involves estimation of the target's pose, which is extracted from the image features in conjunction with its geometric model[27].

The basic component of visual servoing based control scheme is to minimize an error function $\mathbf{e}(t)$. Let us define the kinematic error function as the relative difference between the *pose* vector of the printing nozzle $P_N$ and a stationing point $S_P$ on the deposition platform, both expressed in world frame $\mathcal{F}_\mathcal{C}$ as;

$$\mathbf{e}(t) = P_N - S_P \tag{1}$$

Note that $\hat{\mathbf{S}_P}$ is an estimate of the current position of the stationing point observed in the image frame. Here, $\hat{\mathbf{S}_P} = {}^W x_C . {}^C x_I . {}^I \mathbf{S}_P$. Moreover, $P_N$ can be integrated into two ways, either through the *Jacobian matrix* transformation that relates the indicated values in joint space of the robot to Cartesian coordinates of the end effector. In this case, the system is observing only the target and can be referred to as *endpoint open-loop* (EOL) systems. Equation (1) in the case of *endpoint closed-loop* ECL takes the following form:

$$\mathbf{e}(t) = {}^W x_E \, {}^E P_N - {}^W x_C . {}^C x_I . {}^I S_P \tag{2}$$

Let us consider that the robot task space restricted to $\mathcal{R}^3$ as a rigid link that constraints the pose of the printing nozzle relative to the platform. In this case, the control input the robotic manipulator is desired transitional velocity. To drive the system to an equilibrium state of zero kinematic error, a proportional regulator control is applied:

$$u = -k(\hat{x_E} \, {}^E P_N - \hat{x_C} \, {}^C S_P) \tag{3}$$

Here '$k$' is the $proportional$ feedback constant and is greater than zero. Also, note that Equation (3) is based on EOL system configuration. This can be converted into ECL system by, the visual system observing both the printing nozzle and the platform. In this case Equation (3) becomes:

$$u = -k \, (\hat{x_C} \, {}^C P_N - \hat{x_C} \, {}^C S_P) = -k \, \hat{x_C} \, ({}^C P_N - {}^C S_P) \tag{4}$$

### B. Kalman Filter for Tracking

Kalman filter is an optimal algorithm for estimating target velocity and acceleration at each frame based on the process model describing the motion of the target and camera measurements under noisy conditions. The process model is based on Newton's Equation of motion that can be easily used to predict the future position of the object. In between the measurement, the robotic arm is tracking the position of the target based on the estimated acceleration and velocity. Upon receiving the new observation, the model is updated by taking into account the uncertainty in the measuring and estimation process.

In three dimensions the Newton's Equation of motions can be expressed as a system of equation.

$$\begin{cases} x = x_0 + v_{x0}\Delta t + \frac{1}{2} a_x \Delta t^2 \\ y = y_0 + v_{y0}\Delta t + \frac{1}{2} a_y \Delta t^2 \\ z = z_0 + v_{z0}\Delta t + \frac{1}{2} a_z \Delta t^2 \end{cases} \tag{5}$$

here, the state of the system are the target parameters $[x, y, z, v_x, v_y, v_z, a_x, a_y, a_z]$. The state variables are inputted into the prediction algorithm. Above set of equations describing the relationship between input and output can be referred to as the dynamic or state-space models. Assuming a small time increment and constant acceleration in between, the extrapolated state at time $n$ is described by the following equations of the states.

$$\begin{cases} \hat{x}_{n+1,n} = \hat{x}_{n,n} + \hat{\dot{x}}_{n,n}\Delta t + \frac{1}{2}\hat{\ddot{x}}_{n,n}\Delta t^2 \\ \hat{y}_{n+1,n} = \hat{y}_{n,n} + \hat{\dot{y}}_{n,n}\Delta t + \frac{1}{2}\hat{\ddot{y}}_{n,n}\Delta t^2 \\ \hat{z}_{n+1,n} = \hat{z}_{n,n} + \hat{\dot{z}}_{n,n}\Delta t + \frac{1}{2}\hat{\ddot{z}}_{n,n}\Delta t^2 \\ \hat{\dot{x}}_{n+1,n} = \hat{\dot{x}}_{n,n} + \hat{\ddot{x}}_{n,n}\Delta t \\ \hat{\dot{y}}_{n+1,n} = \hat{\dot{y}}_{n,n} + \hat{\ddot{y}}_{n,n}\Delta t \\ \hat{\dot{z}}_{n+1,n} = \hat{\dot{z}}_{n,n} + \hat{\ddot{z}}_{n,n}\Delta t \\ \hat{\ddot{x}}_{n+1,n} = \hat{\ddot{x}}_{n,n} \\ \hat{\ddot{y}}_{n+1,n} = \hat{\ddot{y}}_{n,n} \\ \hat{\ddot{z}}_{n+1,n} = \hat{\ddot{z}}_{n,n} \end{cases} \tag{6}$$

The general form of the state extrapolation equation in matrix form is [18]

$$\hat{\boldsymbol{x}}_{\boldsymbol{n+1,n}} = \boldsymbol{F}\hat{\boldsymbol{x}}_{\boldsymbol{n,n}} + \boldsymbol{G}\boldsymbol{u}_{\boldsymbol{n}} + \boldsymbol{w}_{\boldsymbol{n}} \tag{7}$$

here, $\hat{x}_{n+1,n}$ is the predicted state of the platform or the tracking object, $\hat{x}_{n,n}$ is the estimated state at time step $n$,$u_n$ is the control or input variable which is an input to the system that is measurable, $w_n$ is the process related noise that not measurable, $F$ is the state transition matrix and $G$ is the control matrix that maps the control to the state variables. To develop an extended Kalman filter for the case of a robotic printer, we consider the control input to be the acceleration vector measured directly from the printing platform. In this case, we have additional information based on the actual movement of

the platform. Next, the generalized observation model in the matrix form is given by:

$$\boldsymbol{z}_n = \boldsymbol{H}\boldsymbol{x}_n + \boldsymbol{v}_n \tag{8}$$

here, $\boldsymbol{z_n}$ is the observation vector, $\boldsymbol{H}$ is the observation matrix, $\boldsymbol{x_n}$ is the system state matrix and $\boldsymbol{v_n}$ is the noise vector related to measurement. Process-related noise matrix Q is evaluated by projecting the variance in acceleration $\sigma^2$ onto the state transition matrix from the dynamic model.

$$\boldsymbol{Q} = \boldsymbol{G}\sigma_a^2\boldsymbol{G^T} \tag{9}$$

the measurement noise is assumed to be described by a zero-mean Gaussian distribution with covariance matrix $\boldsymbol{R}$. It is also assumed that the $x, y, z$ values are uncorrelated.

$$\boldsymbol{R_n} = \begin{bmatrix} \sigma_{x_m}^2 & 0 & 0 \\ 0 & \sigma_{y_m}^2 & 0 \\ 0 & 0 & \sigma_{z_m}^2 \end{bmatrix} \tag{10}$$

here $\sigma_{x_m}^2$ refers to the measurement uncertainty in the $x$ coordinate measurement.

The Kalman filtering process starts with an initial state and covariance matrix $\mathbf{F}$ estimate. A predicted state and uncertainty are computed at each time step through the extrapolation model based on the Equation. (7). Next, the predicted states and uncertainty are corrected based on the received measurements and the associated covariances of the measuring devices. The equations associated with both the steps are summarized below.

$$\begin{gathered} \textit{Prediction Step} \\ \hat{\boldsymbol{x}}_{+1,\boldsymbol{n}} = \boldsymbol{F}\hat{\boldsymbol{x}}_{\boldsymbol{n},\boldsymbol{n}} + \boldsymbol{G}\boldsymbol{u_n} \\ \mathbf{P}_{n+1,n} = \mathbf{F}\mathbf{P}_{n,n}\mathbf{F}^T + \mathbf{Q} \end{gathered} \tag{11}$$

$$\begin{gathered} \textit{Correction Step} \\ \boldsymbol{K_n} = \boldsymbol{P_{n,n-1}}\boldsymbol{H^T}\left(\boldsymbol{H}\boldsymbol{P_{n,n-1}}\boldsymbol{H^T} + \boldsymbol{R}_n\right)^{-1} \\ \hat{\boldsymbol{x}}_{n,n} = \hat{\boldsymbol{x}}_{n,n-1} + \boldsymbol{K}_n\left(z_n - \boldsymbol{H}\hat{\boldsymbol{x}}_{n,n-1}\right) \\ \boldsymbol{P}_{n,n} = \left(\boldsymbol{I} - \boldsymbol{K}_n\boldsymbol{H}\right)\boldsymbol{P}_{n,n-1}\left(\boldsymbol{I} - \boldsymbol{K}_n\boldsymbol{H}\right)^{\boldsymbol{T}} + \boldsymbol{K_n}\boldsymbol{R}_n\boldsymbol{K_n^T} \end{gathered} \tag{12}$$

here $\mathbf{K}_n$ is the Kalman gain at a given time step. Since the initial pose of the target is unknown, an inappropriate estimate can lead to poor convergence of the Kalman filter. To improve the performance, the detected position of the target as retrieved from the camera is used as the initial estimate.

## IV. Results and Conclusion

### A. Experimental Setup

The current experimental testing scenario is as follows; When the platform is stationary, the robotic printer is able to fabricate structures like any other 3D printer. During the printing process, the platform produces uncooperative motion at a uniform velocity. At this time the visual sensor tracks the motions and compensates for this motion in parallel to the printing process. This section presents a detailed description of the experimental setup.

*Positioning System:* The experimental demonstration was carried out using a low-cost, off-the-shelf 6 axis robotic manipulator (Figure (2e)). This product is marketed with a repeat position accuracy of 0.05 mm, rated payload 2 kg, arm length of 650 mm, and a maximum speed when moving in a straight line of 600 mm/s. This arm has an integrated motor controller that can realize numerous motion functions such as linear interpolation and arc interpolation. The motion of the manipulator is controlled by GCODE programming sent from the computer via the serial interface.

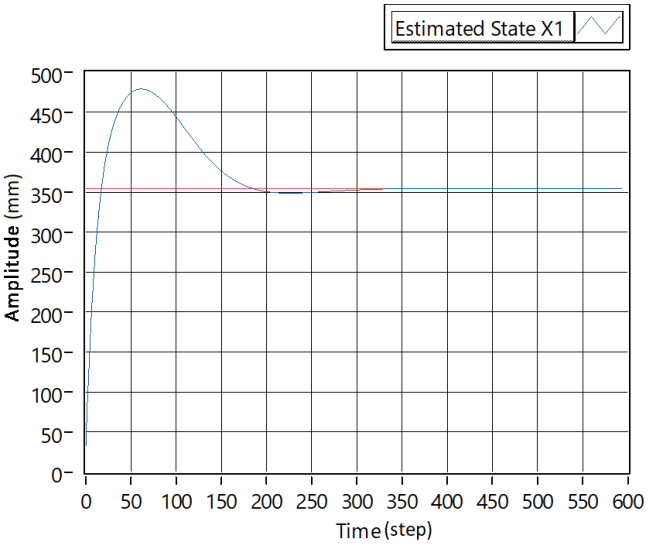


Fig. 1: Step Response of Kalman filter with amplitude 350 mm and time in (ms)

*Visual Tracking:* The visual system comprises a USB camera (Microsoft LifeCam Studio), streaming at 30 $fps$ at a pixel resolution of 1920 x 1080. The camera is calibrated in 2 dimensions to a flat surface which is also the printing surface. Additional inertial measurement (IMU) sensor is attached to the platform and augmented with the Kalman filter in the visual tracker. This IMU sensor provide Cartesian components of acceleration in the platform as the measurable control input variable $u_n$ as described in section. (III-B). This enables the robust performance of the tracking filter in case of highly uncertain and missing measurements. The output from the Kalman filter and the visual sensor input are illustrated in Figure (2a). It is evident that the implemented tracking filter is reliable for smooth trajectory generation of the target from noisy camera inputs.

*Software Platform:* The high-level visual servo controller is deployed as two separate loops running at different step times and sharing information with each other. The inner loop controls the robotic manipulator through GCODE and runs two times faster than the visual loop. The visual loop is restricted with respect to the camera capture rate of 30 fps. Using this approach, the control loop runs independently of the visual loop. This results in two scenarios for the control loop, at each iteration, two of the following scenarios happen, (i) either it is input with new information about the target from the visual cameras, in this case, it updates the current belief of

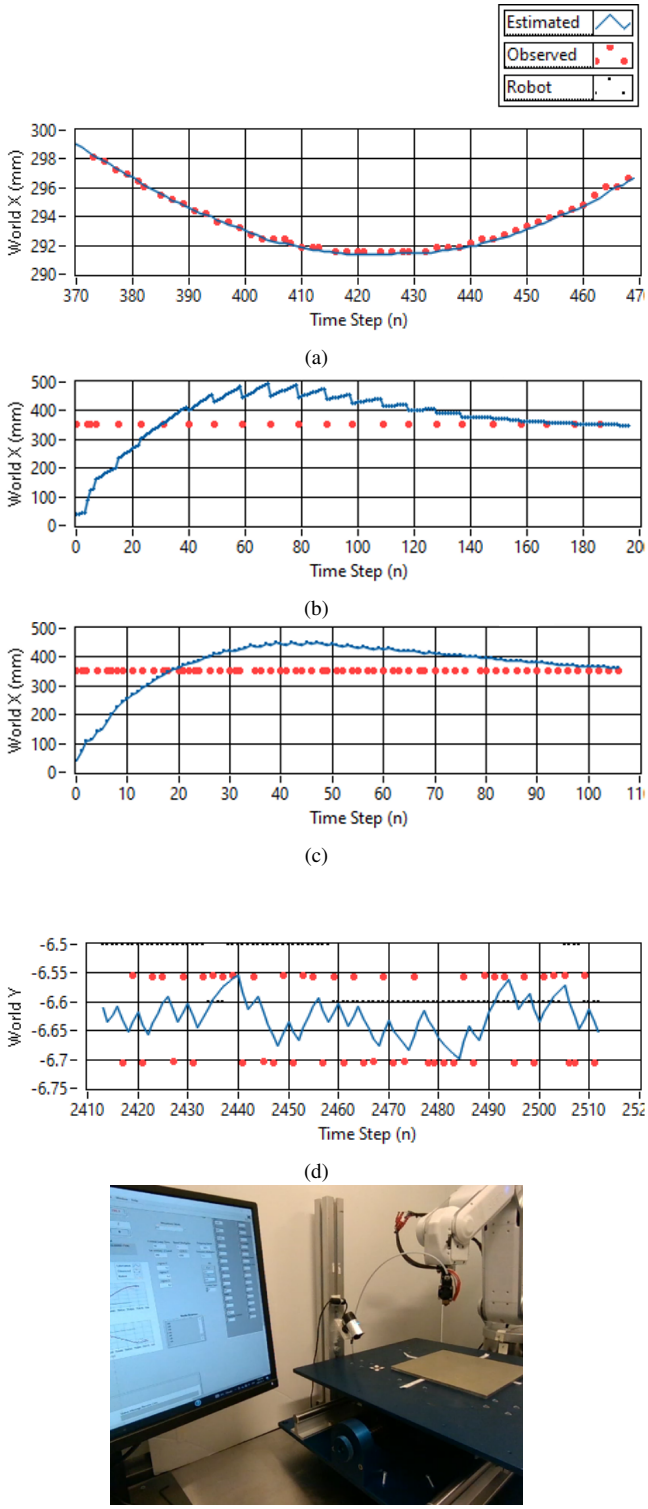


Fig. 2: (a) Visualisation of the state estimation and tracking of a target through Kalman filter from noisy sensor measurements. (b) Step response output with amplitude of 350 mm with $T_{cl} \leq 2(T_{vl})$,(c) Step response output with amplitude of 350 mm with $T_{cl} \leq T_{vl}$, (d) Kalman filter *estimated*, measurement *observed* and robotic state output for a stationary target. Note that the legend is common to all three targets, (e) Overall experimental setup

the target location, (ii) in case no new visual information is available from cameras, the control loop predicts the target location given apriori and system model through the process of Kalman filtering. This can be visualized in Figure (2b and c), which illustrates the step response of 350 mm in two scenarios. In Figure (2b) the control loop is running at least twice as fast as compared to the visual detection loop ( $T_{cl} \leq 2(T_{vl})$). In the case of Figure (2c), the control loop time $T_{cl}$ is equals to the visual detection loop time, $T_{vl}$. It was observed that even when the control loop and visual detection loop run with the time step, the visual detection loop is not able to keep up with the timing as it is waiting on the camera to update the following image. This results in variable loop timing in the visual detection loop. If the visual detection and control loop are not separate, then high-speed control of the robotic manipulator cannot be implemented. Having two parallel loops running and sharing information in real-time can improve the performance of robotic tracking and traversing. The control loop can predict the target's motion.

*Testing Platform:* The platform consists of a flat surface mounted on a servo-controlled linear stage. This platform consists of high accuracy encoders and its positional data can be extracted. This controlled motion is used to validate the calibration of the visual sensors and asses the tracking performance. The platform runs independent of the robotic printing setup and produces motion to simulate an uncooperative platform having 2 degrees of freedom(depending on the pose of the platform relative to the robotic arm). It can generate pulse response and periodic motion. At all times, the camera is tracking a marker located on the printing platform, and this serves as the origin of the platform coordinate frame. When the platform is in motion, the updated printing coordinates are obtained through the beforehand knowledge of platform geometry and the tracked motion.

### B. Conclusion

In this work, a framework for visual detection and control for six dof robotic manipulator is presented. The visual detection and control strategy consists of an inner loop that controls the state of the robot and an outer loop that serves as the visual detection loop. This arrangement enhances the tracking performance as the failure for timely updates from the outer loop is independent of the inner control loop. The overall system is comprised of low-cost, off-shelf ready use components. The Kalman filter enables reliable target state estimation in case of measurement failure and fusion of sensory data. The current performance of this system can be further enhanced by improving the robotic state indications. The system is observed to be lagged by not having timely updates to the robot state due to the manufacturer's controller configuration. In future works, a similar visual detection and tracking of the robotic state itself will be implemented to form the robot's full visual control loop.

## REFERENCES


[1] F. Gharagozloo, B. Tempesta, M. Meyer, D. Nguyen, S. Gruessner, and J. Redan, "History of robotic surgery," in *Robotic Surgery*, pp. 21–29, Springer International Publishing, 2021.

[2] N. Simaan, R. M. Yasin, and L. Wang, "Medical technologies and challenges of robot-assisted minimally invasive intervention and diagnostics," *Annu. Rev. Control Robot. Auton. Syst.*, vol. 1, pp. 465–490, May 2018.

[3] C.-Y. Huang and J. Galeotti, "Robust skin-feature tracking in free-hand video from smartphone or robot-held camera, to enable clinical-tool localization and guidance," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, May 2021.

[4] X. Hu, A. Nguyen, and F. R. y Baena, "Occlusion-robust visual markerless bone tracking for computer-assisted orthopaedic surgery," 2021.

[5] S. Atallah, ed., *Digital Surgery*. Springer International Publishing, 2021.

[6] A. Banach, M. Strydom, A. Jaiprakash, G. Carneiro, A. Eriksson, R. Crawford, and A. McFadyen, "Visual localisation for knee arthroscopy," *Int J CARS*, vol. 16, pp. 2137–2145, July 2021.

[7] L. Li, X. Li, B. Ouyang, S. Ding, S. Yang, and Y. Qu, "Autonomous multiple instruments tracking for robot-assisted laparoscopic surgery with visual tracking space vector method," *IEEE/ASME Trans. Mechatron.*, pp. 1–1, 2021.

[8] M. K. Hasan, L. Calvet, N. Rabbani, and A. Bartoli, "Detection, segmentation, and 3d pose estimation of surgical tools using convolutional neural networks and algebraic geometry," *Medical Image Analysis*, vol. 70, p. 101994, May 2021.

[9] S. Speidel, E. Kuhn, S. Bodenstedt, S. Röhl, H. Kenngott, B. Müller-Stich, and R. Dillmann, "Visual tracking of da vinci instruments for laparoscopic surgery," in *SPIE Proceedings*, SPIE, Mar. 2014.

[10] I. Laina, N. Rieke, C. Rupprecht, J. P. Vizcaíno, A. Eslami, F. Tombari, and N. Navab, "Concurrent segmentation and localization for tracking of surgical instruments," in *Lecture Notes in Computer Science*, pp. 664–672, Springer International Publishing, 2017.

[11] R. Moccia, C. Iacono, B. Siciliano, and F. Ficuciello, "Vision-based dynamic virtual fixtures for tools collision avoidance in robotic surgery," *IEEE Robot. Autom. Lett.*, vol. 5, pp. 1650–1655, Apr. 2020.

[12] D. Bouget, M. Allan, D. Stoyanov, and P. Jannin, "Vision-based and marker-less surgical tool detection and tracking: a review of the literature," *Medical Image Analysis*, vol. 35, pp. 633–654, Jan. 2017.

[13] Z. Zhao, S. Voros, Y. Weng, F. Chang, and R. Li, "Tracking-by-detection of surgical instruments in minimally invasive surgery via the convolutional neural network deep learning-based method," *Computer Assisted Surgery*, vol. 22, pp. 26–35, Sept. 2017.

[14] F. Qin, Y. Li, Y.-H. Su, D. Xu, and B. Hannaford, "Surgical instrument segmentation for endoscopic vision with data fusion of cnn prediction and kinematic pose," in *2019 International Conference on Robotics and Automation (ICRA)*, IEEE, May 2019.

[15] T. Kurmann, P. Marquez Neila, X. Du, P. Fua, D. Stoyanov, S. Wolf, and R. Sznitman, "Simultaneous recognition and pose estimation of instruments in minimally invasive surgery," in *Lecture Notes in Computer Science*, pp. 505–513, Springer International Publishing, 2017.

[16] S. Giannarou, M. Visentini-Scarzanella, and G.-Z. Yang, "Probabilistic tracking of affine-invariant anisotropic regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, pp. 130–143, Jan. 2013.

[17] A. Schoob, D. Kundrat, L. A. Kahrs, and T. Ortmaier, "Stereo vision-based tracking of soft tissue motion with application to online ablation control in laser microsurgery," *Medical Image Analysis*, vol. 40, pp. 80–95, Aug. 2017.

[18] H. Chen, W. Zhang, and D. Yan, "Robust visual tracking with reliable object information and kalman filter," *Sensors*, vol. 21, p. 889, Jan. 2021.

[19] J. Kummert, A. Schulz, T. Redick, N. Ayoub, A. Modabber, D. Abel, and B. Hammer, "Efficient reject options for particle filter object tracking in medical applications," *Sensors*, vol. 21, p. 2114, Mar. 2021.

[20] F. Vasconcelos, E. Mazomenos, J. Kelly, and D. Stoyanov, "Rcm-slam: Visual localisation and mapping under remote centre of motion constraints," in *2019 International Conference on Robotics and Automation (ICRA)*, IEEE, May 2019.

[21] S. Bi, Y. Gu, Z. Zhang, H. Liu, C. Zhai, and M. Gong, "Multi-camera stereo vision based on weights," in *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, IEEE, May 2020.

[22] X. Du, M. Allan, A. Dore, S. Ourselin, D. Hawkes, J. D. Kelly, and D. Stoyanov, "Combined 2d and 3d tracking of surgical instruments for minimally invasive and robotic-assisted surgery," *Int J CARS*, vol. 11, pp. 1109–1119, Apr. 2016.

[23] N. Rieke, D. J. Tan, C. Amat di San Filippo, F. Tombari, M. Alsheakhali, V. Belagiannis, A. Eslami, and N. Navab, "Real-time localization of articulated surgical instruments in retinal microsurgery," *Medical Image Analysis*, vol. 34, pp. 82–100, Dec. 2016.

[24] M. Allan, S. Ourselin, D. J. Hawkes, J. D. Kelly, and D. Stoyanov, "3-d pose estimation of articulated instruments in robotic minimally invasive surgery," *IEEE Trans. Med. Imaging*, vol. 37, pp. 1204–1213, May 2018.

[25] X. Du, N. Clancy, S. Arya, G. B. Hanna, J. Kelly, D. S. Elson, and D. Stoyanov, "Robust surface tracking combining features, intensity and illumination compensation," *Int J CARS*, vol. 10, pp. 1915–1926, June 2015.

[26] B. Yang, W.-K. Wong, C. Liu, and P. Poignet, "3d soft-tissue tracking using spatial-color joint probability distribution and thin-plate spline model," *Pattern Recognition*, vol. 47, pp. 2962–2973, Sept. 2014.

[27] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Automat.*, vol. 12, no. 5, pp. 651–670, 1996.